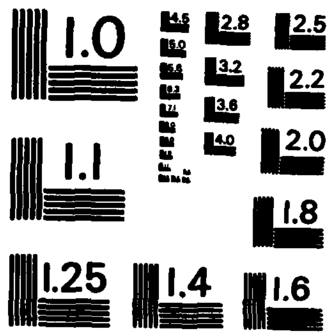
AD-A160 177 AN EFFICIENT RLS (RECURSIVE-LEAST-SQUARES) DATA-DRIVEN ECHO CANCELLER FOR..(U) STANFORD UNIV CA INFORMATION SYSTEMS LAB J M CIOFFI ET AL. JUN 85 AFOSR-TR-85-0762 1/1

UNCLASSIFIED F49620-79-C-0058 F/G 17/2 NL

END

FILMED

DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

(2)

REPORT DOCUMENTATION PAGE

AD-A160 177

1b. RESTRICTIVE MARKINGS	
2b. DECLASSIFICATION/DOWNGRADING SCHEDULE: N/A	3. DISTRIBUTION/AVAILABILITY OF REPORT: Approved for public release; distribution unlimited
4. PERFORMING ORGANIZATION REPORT NUMBER(S)	5. MONITORING ORGANIZATION REPORT NUMBER(S): AFOSR-TR- 85-0762
6a. NAME OF PERFORMING ORGANIZATION: Stanford University / 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION: AFOSR
6c. ADDRESS (City, State and ZIP Code): Department of Electrical Engineering, Stanford, CA 94305	7b. ADDRESS (City, State and ZIP Code): Bldg. 410, Bolling AFB, D.C. 20332-6448
8a. NAME OF FUNDING/SPONSORING ORGANIZATION: AFOSR / 8b. OFFICE SYMBOL (If applicable): NM	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER: F49620-79-C-0058
8c. ADDRESS (City, State and ZIP Code): Bldg. 410, Bolling AFB, D.C. 20332-6448	10. SOURCE OF FUNDING NOS.

PROGRAM ELEMENT NO.	PROJECT NO.	TASK NO.	WORK UNIT NO.
61102F	2304	A6	

11. TITLE (Include Security Classification) An Efficient, R.L.S. Data-Driven Echo Canceller for Fast Initialization of Full-Duplex Data Transmission

12. PERSONAL AUTHOR(S) J. M. Cioffi, T. Kailath

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Yr., Mo., Day)	15. PAGE COUNT
Reprint	FROM ______ TO ______	June 1985	5

16. SUPPLEMENTARY NOTATION

DTIC ELECTE OCT 15 1985 S E

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB. GR.	RLS, echo cancellor, initializition, full duplex data fast transversal filter
XXXXX	XXXXXXXXX	XXX	

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

Computationally efficient Recursive-Least-Squares (RLS) procedures are presented specifically for the adaptive adjustment of the Data-Driven Echo Cancellers (DDECs) that are used in voiceband full-duplex data transmission. The methods are shown to yield very short learning times for the DDEC while they also simultaneouwly reduce computational requirements to below those required for other least-square procedures, such as those recently proposed by Salz (1983). The new methods can be used with any training sequence over any number of iterations, unlike any of the previous fast-converging methods. Tbe methods are based upon the Fast Transversal Filter (FTF) RLS adaptive filtering algorithms that were independently introduced by the authors of this paper; however, several special features of the DDEC are introduced and exploited to further reduce computation to the levels that would be required for slower-converging stochastic-gradient solutions. Several trade-offs between computation, memory, learning-time and performance are also illuminated for the new initialization.

20. DISTRIBUTION/AVAILABILITY OF ABSTRACT	21. ABSTRACT SECURITY CLASSIFICATION
UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☒ DTIC USERS ☐	Unclassified

22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE NUMBER (Include Area Code)	22c. OFFICE SYMBOL
Dr. Marc Q. Jacobs	(202)767-4940	NM

DTIC FILE COPY

DD FORM 1473, 83 APR EDITION OF 1 JAN 73 IS OBSOLETE.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE

AN EFFICIENT, RLS, DATA-DRIVEN ECHO CANCELLER FOR FAST INITIALIZATION OF FULL-DUPLEX DATA TRANSMISSION

John M. Cioffi
IBM Research Laboratory
San Jose, CA 95193

T. Kailath
Information Systems Laboratory
Stanford, CA 94305

ABSTRACT

Computationally efficient Recursive-Least-Squares (RLS) procedures are presented specifically for the adaptive adjustment of the Data-Driven Echo Cancellers (DDECs) that are used in voiceband full-duplex data transmission. The methods are shown to yield very short learning times for the DDEC while they also simultaneously reduce computational requirements to below those required for other least-square procedures, such as those recently proposed by Salz (1983). The new methods can be used with any training sequence over any number of iterations, unlike any of the previous fast-converging methods. The methods are based upon the Fast Transversal Filter (FTF) RLS adaptive filtering algorithms that were independently introduced by the authors of this paper; however, several special features of the DDEC are introduced and exploited to further reduce computation to the levels that would be required for slower-converging stochastic-gradient solutions. Several trade-offs between computation, memory, learning-time, and performance are also illuminated for the new initialization.

1. INTRODUCTION

Echo cancellers were suggested for use in 2-wire full-duplex data transmission by Koll and Weinstein [1] in 1973. Additionally, much other work concerning the data echo canceller has appeared in [2-12,19]. Of particular concern in this paper is Mueller's Data-Driven Echo Canceller (DDEC) [6]. The use of stochastic-gradient LMS (Least-Mean-Square) algorithms in the DDEC has led to unacceptably long training periods for the full-duplex modem [2,11,12]. In this paper, we specifically investigate the use of the Fast-Transversal-Filters (FTF) Recursive-Least-Squares (RLS) adaptive algorithms [13-14] in the DDEC to substantially reduce the necessary training period.

Because the transmitted data sequence is usually "whitened" through scrambling prior to entering the transmitter and DDEC, it was orginally believed that the use of RLS adaptive algorithms would have led to no improvement in the convergence time of the DDEC in comparison to stochastic-gradient techniques. However, Farrow [15], Honig [11] and Salz [12] verified a significant convergence improvement (about a factor of 5, see [11]) of the RLS (or closely related) methods when the double-taking data signal was intentionally silenced during the initial training phase of the data transmission. However, there are several drawbacks of the "Salz-Farrow" (SF) method in [11], [12], and [15]. Most of these are the result of the SF method's absolute necessity for the training sequence to be "pseudorandom" with very special autocorrelation properties and with a period equal to the number of coefficients (order) of the echo canceller, which limits both the permissible orders (to, say, 7, 15, 31, 63, 127, 255, 511, ..., 2^n-1) and the performance of the RLS DDEC.

This paper introduces FTF solutions that require less computation than the SF method, permit training of the echo canceller with any known training sequence of any length (long enough to converge), and which converge as fast or faster than the SF method. The freedom of choice in training sequence can also result in a factor 2 or more improvement (reduction) in learning time to get to the same echo canceller performance level as the SF echo canceller. We specifically investigate many interesting trade-offs between computational requirements and the performance of the echo canceller.

The new method's much-less-restrictive or arbitrary choice of training sequence permits use of sequences that are "white" (autocorrelation matrix is a diagonal), such as those of Milewski [17]. The use of such a sequence can lead to as much as a 3dB advantage over the pseudorandom sequence. Also, the prewindowed FTF solutions do not require "priming" the echo channel with N inputs, before computation can begin, as is necessary in the SF method, and are numerically stable over the initialization time period (using the soft-constraint initialization of [13,14]).

A possible disavantage, however, of the new method is its requirement of more read-only memory than the SF method, if one wishes to keep computation to an absolute minimum. This extra memory is used to pre-store certain quantities of the RLS algorithms that are solely a function of the known training sequence. Since practical experience dictates that the cost of read-only memory, in comparison to the cost of the other signal-processing functions that appear in high-speed modems, is low; this possible disadvantage is minimal.

Section 2 reviews and analyzes the RLS DDEC. Section 3 introduces and discusses the new recursive initialization procedures. Finally, Section 4 is a brief conclusion. A longer version of this paper appears in [16].

2. RLS AND THE DDEC

This section briefly reviews the DDEC and the application of RLS methods to it.

2.1 Definitions and Terminology

The near-end transmitted data signal is defined as

$$u_1(t) \triangleq \mathrm{Re}\left\{\sum x(kT_s)p(t-kT_s)e^{j\omega_c t}\right\}. \qquad 2\text{-}1$$

where the inphase and quadrature data symbols are the real and imaginary parts of $x(kT_s)$, respectively. The carrier frequency is $\omega_c/2\pi$, the baseband pulse shaping is $p(t)$, and $1/T_s$ is the symbol rate. Also, "Re" denotes the real part of a complex number. $u_1(t)$ is the real part of analytic signal, $U_1(t)$,

$$U_1(t) = \sum_k x(kT_s)p(t-kT_s)e^{j\omega_c t}. \qquad 2\text{-}2$$

The impulse response of the combined hybrid and channel path is $h(t)$. The hybrid output, $d(t)$, is the sum of the echo, and the uncorrelated double-talking data signal and channel noise $u_2(t)$,

$$d(t) = h(t) * u_1(t) + u_2(t), \qquad 2\text{-}3$$

where * denotes continuous-time convolution.

$h(t)$ is also written as

$$h(t) = \mathrm{Re}\left\{h_{BB}(t)e^{j\omega_c t}\right\}, \qquad 2\text{-}4$$

where $h_{BB}(t)$ is the baseband-equivalent [18] echo path for $h(t)$.

This work was supported in part by the U.S. Army Research Office, under Contract DAAG29-79-C-0215, and by the Air Force Office of Scientific Research, Air Force Systems Command, under Contract AF49-620-79-C-0058.

Then d(t) becomes (see [18])

$$d(t) = \mathrm{Re}\left\{\sum_k x(kT_s)g(t-kT_s)e^{j\omega_c t}\right\} + n_2(t), \qquad 2\text{-}5$$

where g(t) is

$$g(t) = \frac{1}{2}\,p(t) * h_{BB}(t). \qquad 2\text{-}6$$

The echo estimate is $\hat{d}(t)$, and the error signal (double-talker estimate) is

$$\epsilon(t) = d(t) - \hat{d}(t). \qquad 2\text{-}7$$

The Minimum Mean-Square-Error (MMSE) of $\epsilon(t)$ is

$$\mathrm{MMSE} \triangleq \min_{\hat{d}} E\left[\epsilon(t)^2\right] \triangleq \sigma_2^2, \qquad 2\text{-}8$$

where $\hat{d}(t)$ becomes

$$\hat{d}(t) = \mathrm{Re}\left\{\sum_k x(kT_s)g(t-kT_s)e^{j\omega_c t}\right\}, \qquad 2\text{-}9$$

assuming a linear-time-invariant estimator followed by a modulator. $E[\cdot]$ denotes expectation.

2.2 The Data-Driven Echo Canceller (DDEC)

The DDEC appears in Figure 1. The adaptive transversal filter acts at a tap-spacing, T_s/l, that is sufficiently short to satisfy the Sampling Theorem for the entire passband transmitted signal $u_1(t)$, where l is integer. The continuous $U_1(t)$ is rewritten

$$U_1(t) = \sum_k x(kT_s)e^{j\omega_c kT_s}g(t-kT_s)e^{j\omega_c(t-kT_s)} \qquad 2\text{-}10a$$

$$= \sum_k \tilde{x}(kT_s)\tilde{g}(t-kT_s) \qquad 2\text{-}10b$$

$$\text{where } \tilde{x}(kT_s) = x(kT_s)e^{j\omega_c kT_s} \qquad 2\text{-}11$$

$$\text{and } \tilde{g}(t) = g(t)e^{j\omega_c t}, \qquad 2\text{-}12$$

as discussed in [2]. In conventional echo-cancellation schemes [2], the designer carefully chooses ω_c and T_s so that the rotation of the data symbols in Equation 2-11 is trivial (typically 270°). Thus, we now drop the tilde on $\tilde{x}(kT_s) \rightarrow x(kT_s)$ in the ensuing results. The DDEC then synthesizes $\tilde{g}(t)$ at rate T_s/l

Since d(t) is real, $\hat{d}(t)$ should also be real, or

$$\hat{d}(kT_s) = W^R_{N,k}\mathrm{Re}\{\bar{X}_N(k)\} - W^Q_{N,k}\mathrm{Im}\{\bar{X}_N(k)\} \qquad 2\text{-}13$$

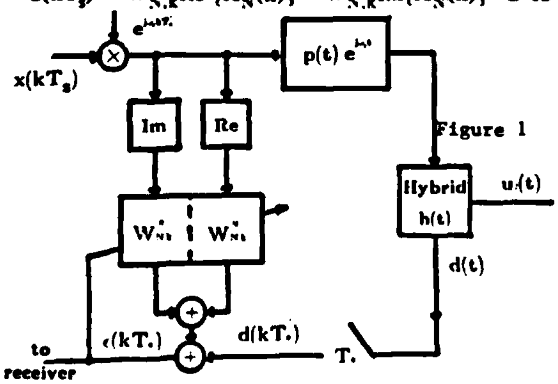


Figure 1

where "Im" denotes imaginary part, $W^R_{N,k}$ and $W^Q_{N,k}$ are the real and imaginary parts of the adaptive transversal filter (complex $l \times Nl$ row vector) that estimates $\tilde{g}(t)$, N is the order of (or number of spanned symbol periods in) the DDEC, and $X_N(k)$ is the column vector ($Nl \times 1$) corresponding to the last N DDEC inputs at the sampling rate $1/T_s$.

$$\bar{X}_N(k) \triangleq \left[x(kT_s)\ 0\ \dots\ 0\ \mid\ x([k-N+1]T_s)\ 0\ \dots\ 0\right]^T \qquad 2\text{-}14$$

where a superscript of T denotes transpose.

2.3 Subcancellers

Detailed analyses of the use of subcancellers appear in [2,3,5]. The essential structural simplification arises because N (not Nl) taps in the transversal filters contribute to d(T) at any sampling instant (see zeros in 2-14). The structure is equivalent to l sub-echo cancellers or "subcancellers" that independently act to estimate the l phases (per symbol period) of the desired echo-contaminated output. We add the new observation that the same inputs appear in each subcanceller, and the majority of computation in the FTF (or any fast-RLS) algorithms, which depends only on these inputs, need only be performed once for the group of subcancellers, even when the training sequence is unknown. Since $l \geq 4$ in practical voiceband modems for full-duplex data communications, this leads to large computational and storage savings.

2.4 The Application of RLS to the DDEC

The RLS DDEC chooses $W^i_{N,k} = W^{R_i}_{N,k} + jW^{Q_i}_{N,k}$ to minimize the (for the i^{th} subcanceller)

$$\epsilon^i_N(k) = \sum_{m=0}^{k}\left(d(mT_s + iT_s') + W^{R_i}_{N,k}X^R_N(mT_s) + W^{Q_i}_{N,k}X^Q_N(mT_s)\right)^2, \qquad 2\text{-}15$$

where $i=0, \dots, l-1$, $X^R_N(mT_s)$ and $X^Q_N(mT_s)$ are the real and imaginary parts of

$$X_N(mT_s) \triangleq \begin{matrix} x(mT_s) \\ x\left([m-N+1]T_s\right) \end{matrix}, \qquad 2\text{-}16$$

respectively, and $T_s' \triangleq T_s/l$. Minimization of Equation 2-15 directly requires the two-channel (real) FTF algorithms of [13-14]. However, a single-channel FTF algorithm can be used (during training) by staggering the inphase and quadrature data sequences (setting either or both to zero at appropriate time instants). The single-channel is less costly to implement than the two-channel. More about specific implementational comparisons appear in Section 3.

The solution to 2-15, when the DDEC staggers the starting sequence, is

$$W^{R_i}_{N,k} = \left(\sum_{m=0}^{k} d(mT_s + iT_s')X^R_N(mT_s)\right)^T \cdot \qquad 2\text{-}17a$$

$$\left(\sum_{m=0}^{k} X^R_N(mT_s)X^R_N(mT_s)^T\right)^{-1} \text{ and}$$

$$W^{Q_i}_{N,k} = \left(\sum_{m=0}^{k} d[iT_s' + (m+k+N)T_s]X^Q_N[(m+k+N)T_s]\right)^T \qquad 2\text{-}17b$$

$$\left(\sum_{m=0}^{k} X^Q_N[(m+k+N)T_s]X^Q_N[(m+k+N)T_s]^T\right)^{-1}$$

where $k \geq N$, and k is the number of learning iterations for the subcanceller. (The reason for k+N in Equation 2-17b is

discussed in Section 3, and has to do with the aforementioned staggering.) No generality or performance is lost if one by choosing $X_N(mT_s)$ such that

$$X_N^Q(mT_s + (k + N)T_s) = X_N^R(mT_s) \quad 0 \le m \le k-1 . \quad 2\text{-}18$$

Thus, for all $2l$ subcancellers (l inphase and l quadrature), one need only invert the same matrix

$$R_N(k) \triangleq \sum_{m=0}^{k} X_N^R(mT_s) X_N^R(mT_s)' . \quad 2\text{-}19$$

The FTF methods invert this matrix only implicity to obtain an equivalent set (much less storage) of parameters ($C_{N,k}$ filters or "Kalman Gain," see [16]). All of the computation in the equivalent of the inversions can be off line since the training sequence $X_N(kT_s)$ is known beforehand. No useful off-line computation can be performed in the SF method, which also is restricted to the use of pseudorandom training sequences.

2.5 Performance Analysis of RLS DDEC

We assume that NT_s exceeds the nonzero time extent of $\tilde{g}(t)$. Under this common assumption, the RLS estimates of $W_{N,k}^R$ and $W_{N,k}^Q$ are unbiased after $k \ge N$ iterations (see [13,14]), that is

$$E[W_{N,k}^R] = \mathrm{Re}\{[\tilde{g}(iT_s')...\tilde{g}((k - N + 1)T_s + iT_s'))]\} \quad 2\text{-}20a$$

$$E[W_{N,k}^Q] = \mathrm{Im}\{\tilde{G}_i\} , \quad 2\text{-}20b$$

irrespective of the double-taker's presence. G_i is the impulse response for the i^{th} echo subchannel. This unbiased estimator property is not exhibited by stochastic-gradient solutions until a much later time. The (RLS) covariance matrix ([13,14]) for either $[W_{N,k}^R$ or $W_{N,k}^Q]$ is (for $k \ge N$, and white channel noise)

$$\mathrm{cov}[W_{N,k}^{R_i}] = \mathrm{cov}[W_{N,k}^{Q_i}] = R_N^{-1}(k)\sigma_2^2 , \quad 2\text{-}21$$

$$\text{where } \sigma_2^2 \triangleq E[u_2(t)^2] , \quad 2\text{-}22$$

and $R_N(k)$ is given in 2-19. Thus, the RLS solution is near optimum after N iterations only if σ_2^2 is small. One achieves small σ_2^2 by intentionally silencing the double-talking data signal during training [12]. This leaves σ_2^2 equal to the residual channel-noise power level, which is typically much smaller than the power levels of the other signals in the problem. However, the stochastic-gradient methods will still be far from optimum because of the biased mean of $W_{N,k}^R$ or $W_{N,k}^Q$ after only N iterations. They take about 5-10 times longer [11,16].

Figure 2 simulates a situation typical of 4800 bps full-duplex data transmission. The order is 22, while the number of subcancellers l is 4. The performance improvement of the RLS methods is illustrated by the staggered prewindowed FTF solution, which permits training more rapidly than the stochastic-gradient solution.

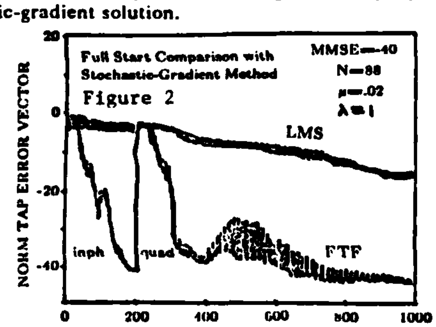


Figure 2

We now turn to comparing the various RLS solutions. Paralleling Salz [12], we can use the trace of cov $[W_{N,k}^R]$ as an indicator of equality. Salz [12] shows that if a length-N pseudorandom sequence is used, then

$$\mathrm{trace}\left\{\mathrm{cov}\left[W_{N,N-1}^{R_i}\right]\right\} = \frac{2N}{N+1}\sigma_2^2 , \quad 2\text{-}23$$

while it is trival to show for a pure white training sequence (after N iterations)

$$\mathrm{trace}\left\{\mathrm{cov}\left[W_{N,N-1}^{R_i}\right]\right\} = \sigma_2^2 . \quad 2\text{-}24$$

Thus, the pseudorandom training sequence is about 3dB worse over N iterations than a white training sequence, under the $\mathrm{cov}[W_{N,k}^R]$ criteria. Equivalently, it is also easy to show [12] (for pseudorandom) that over 2N iterations

$$\mathrm{trace}\left\{\mathrm{cov}\left[W_{N,2N-1}^{R_i}\right]\right\} = \frac{N}{N+1}\sigma_2^2 , \quad 2\text{-}25$$

or it takes about twice as long to reach the same performance level with pseudorandom training in comparison to white training.

The above performance comparisons also hold for the multichannel case, see [16]. responses are simultaneously computed.

We also show in Section 3, after further defining the windowing methods, that in terms of double-talker estimate quality, the pseudorandom starting sequence is the worst possible choice, so that the FTF methods offer substantial improvements in comparison to the SF methods.

3. RLS DDEC ALGORITHM COMPARISON

This section lists and compares the various initialization procedures for the RLS DDEC. (See Tables 1-3).

3.1 New FTF Solutions for DDEC Initialization

An important component in assessing performance and learning time of the RLS DDEC is the data window for the sum-of-squares-errors criterion (equation 2-15). In the DDEC, essentially two windowing cases are of interest: the prewindowed case and the Growing-Memory Covariance (GMC-"unwindowed") case (see [13,14]). The prewindowed FTF solutions assume that all data before the very first iteration is zero. The more general GMC case allows this data to take arbitrary values. The GMC method is only necessary for the DDEC if one desires the autocorrelation matrix to assume some exact, prespecified form on the N^{th} iteration of the initialization, such as is in the SF method [12]. This fixing of the autocorrelation matrix mandates the "priming" of the echo channel with approximately N nonzero data values prior to the first iteration of the algorithm, which adds an additional N (2N in multichannel or QAM case) delay (in symbol periods) to the learning time. In the prewindowed solution, there is absolutely no need for this priming, thus leading to a reduction in learning time. Both experimentally and analytically, the elimination of priming is not a significant drawback for the prewindowed algorithm.

Another important component, in terms of learning time and computation, of the DDEC intialization is the choice of a single-channel or a multichannel solution. The staggered single-channel solution requires one-half the computation of the multichannel solution, but can lead to an extra N units of delay in the prewindowed case. Specifically, the proposed staggered single-channel solution first transmits and trains upon the inphase echo channel ($W_{N,T}^{R_i}$ $i = 0, ..., l - 1$), while simultaneously zeroing (suppressing) the quadrature training sequence. Then N symbol periods of suppressing both inphase and quadrature sequences follow to clear the echo channel. The third and final step is to now transmit only the quadrature training sequence (usually the same sequence, see Equation 2-17b), while suppressing inphase signals. Since both inphase

and quadrature estimation is separate in time, a single channel algorithm is used twice, once for inphase training, and once for quadrature training.

The prewindowed single-channel (staggered) FTF algorithm is then ($k \geq N-1$)

Prewindowed (Single-Channel) FTF ($i = 0, ..., l-1$)

i). $0 \leq m \leq k$ (zero quadrature training sequence)

$$\epsilon_N^{R_i}(m) = d(mT_s + iT'_s) + W^{R_i}_{N,m-1} X_N^R(mT_s) \quad \text{3-1a}$$

$$W^{R_i}_{N,m} = W^{R_i}_{N,m-1} + \epsilon_N^{R_i}(m) C_{N,m} \quad \text{3-1b}$$

ii). $k < m \leq k+N-1$ (zero both inphase and quadrature training sequences)

iii). $k+N-1 < m \leq 2k+N-1$ (zero inphase training sequence)

$$\epsilon_N^{Q_i}(m) = d(\tilde{m}T_s + iT'_s) + W^{Q_i}_{N,m-1} X_N^Q(mT_s) \quad \text{3-1c}$$

$$W^{Q_i}_{N,m} = W^{Q_i}_{N,m-1} + \epsilon_N^{Q_i}(m-1) C_{N,m}\,, \quad \text{3-1d}$$

while the GMC (unwindowed) version is

Covariance (Single-Channel) FTF ($i = 0, ..., l-1$)

i). $0 \leq m \leq N-2$ (nonzero inphase priming, zero quadrature)

$$\epsilon_N^{R_i}(m) = d\left(mT_s + iT'_s\right) + W^{R_i}_{N,m-1} X_N^R(mT_s) \quad \text{3-2a}$$

$$W^{R_i}_{N,m} = W^{R_i}_{N,m-1} + \epsilon_N^{R_i}(m) C_{N,m} \quad \text{3-2b}$$

iii). $k+N \leq m \leq k+2N-1$ (zero inphase, prime quadrature)

iv). $k+2N \leq m \leq 2k+2N$ (while zeroing inphase)

$$\epsilon_N^{Q_i}(m) = d\left(mT_s + iT'_s\right) + W^{Q_i}_{N,m} X_N^Q(mT_s) \quad \text{3-2c}$$

$$W^{Q_i}_{N,m} = W^{Q_i}_{N,m-1} + \epsilon_N^{Q_i}(m) C_{N,m}\,. \quad \text{3-2d}$$

The filter $C_{N,m}$ is computed from the known training sequence beforehand and stored for $0 \leq m \leq k$, (see [16]). In general in the prewindowed initialization, at least one-half of the total coefficients (of $C_{N,m}$) are always zero, leading to a reduction in both computation and storage in comparison to the covariance case, in which no such simplification generally arises. The covariance algorithm can use a training sequence that is exactly white over N iterations.

The multichannel algorithms determine the inphase and quadrature responses simultaneously. The multichannel prewindowed algorithm is ($k \geq 2N-1$)

Prewindowed (Multichannel) FTF ($i=0, ..., l-1$)

$0 \leq m \leq k$

$$\epsilon_N^{RQ_i}(m) = d(mT_s + iT'_s) + W^{RQ_i}_{N,m-1} X_N^{RQ}(mT_s) \quad \text{3-3a}$$

$$W^{RQ_i}_{N,m} = W^{RQ_i}_{N,m-1} + \epsilon_N^{RQ_i}(m) C_{N,m}\,, \quad \text{3-3b}$$

while the covariance case is

Covariance (Multichannel) FTF ($i=0, ..., l-1$)

i). $0 \leq m \leq 2N-1$ (prime inphase and quadrature)

ii). $2N \leq m \leq k+2N$

$$\epsilon_N^{RQ_i}(m) = d\left(mT_s + iT'_s\right) + W^{RQ_i}_{N,m-1} X_N^{RQ}(mT_s) \quad \text{3-4a}$$

$$W^{RQ_i}_{N,m} = W^{RQ_i}_{N,m-1} + \epsilon_N^{RQ_i}(m) C_{N,m}\,. \quad \text{3-4b}$$

Table 1 compares the algorithms in Equations 3-1 through 3-4.

In Table 1, the single-channel (staggered) prewindowed FTF (Equations 3-1a, b, c, d) has the lowest computational requirements. The operation of this particular method was verified in Figure 2. Table 2 is a specific comparison for the conditions of figure 2.

Table 1
DDEC INITIALIZATION COMPARISON

Algorithm	Computation/Iteration	# of active iterations	Total Computation	Extra Delay	Start-up Time
Covariance (Unwindowed) Solutions					
FTF Multichannel Single-channel	4N 2N	2N 2N	$8N^2$ $4N^2$	2N at beginning N begin / N middle	4N 4N
SIF Multichannel Single-channel	4N 2N	4N 4N	$16N^2$ $8N^2$	2N at beginning N begin / N middle	4N 4N
Prewindowed Solutions					
FTF Multichannel Single-channel	2N N	2N 2N	$4N^2$ $2N^2$	0 N middle	2N 3N
Exact FTF Unstable (large N)	N/2	2N	N^2	N middle	2N

In Figure 2, we chose the starting sequence arbitrarily (actually pseudorandom sequence of length >65,000>>22). The true echo-channel impulse response was known for the simulation, and we computed and plotted the quantity of Equations 2-23 and 2-24 (norm tap error vector). There, one sees that the DDEC converges in about $k=N(=22)$ iterations at T_s (or 88 iterations at T_s as is shown in Figure 2), and the choice of training sequence is not critical. We have used $k=100$ iterations at rate T_s (and 100 iterations of channel clearing between inphase and quadrature) to illustrate another advantage of our approach that the FTF solutions can be propagated for any number of iterations to further "fine-tune" the solution. However, the minimum of 2N is used for the total prewindowed learning time and total computation figures in Table 1. The short learning period in all the methods of Table 1 is caused by σ_2^2 being very low (double-talker is silenced for training).

Furthermore, one uses the formula for excess error from Section 2.2 of [13] to obtain

$$\text{excess MSE} = \sigma_2^2(1 - \gamma_N(k)) \quad \text{3-5}$$

where

$$\gamma_N(k) \triangleq 1 - X_N^R(k) R_N^{-1}(k) X_N^R(k)\,, \quad \text{3-6}$$

The worst (maximum of 3-5) that the echo canceller can do at any time ($k \geq N-1$) is

$$\text{excess MSE} = \sigma_2^2\,. \quad \text{3-7}$$

In this case, the worst possible RLS MSE after echo cancellation is thus

$$\text{MSE} = 2\sigma_2^2\,. \quad \text{3-8}$$

This worst possible performance of prewindowed RLS (which is nevertheless a dramatic improvement over stochastic-gradient methods) is achieved by the pseudorandom-trained SF method or the exactly white-trained GMC FTF method when T=N. Thus, any other training sequence† for the FTF performs at least as well under the excess MSE measure.

3.2 The SF Method

[16] lists the SF method in terms of the quantities defined in this paper. One should note immediately that in only the SF methods is the number of iterations frozen beforehand. Table 1 lists the SF methods as covariance methods, since they require priming of the channel with the pseudorandom sequence before the first iteration to ensure the desired structure of the underlying autocorrelation matrix.

3.3 Storage Requirements (Initialization)

†The training sequence is not completely arbitrary in that the autocorrelation matrix must be nonsingular, precluding ridiculous choices such as all zeros or DC.

The random-access-memory (RAM) requirements of all the various algorithms above are about the same, 2Nl+N RAM storage locations. However, the proposed FTF methods of this chapter also require a significant amount of read only memory (ROM) if one desires to store the quantity $C_{N,m}$ over the initialization interval rather than compute it on line. Just how much storage depends upon the window and also upon the number of channels (single- or multichannel). The storage requirements appear in Table 3, in general and under the conditions of Figure 3 (N=22, l=4). There one determines that the ROM requirements are not substantial by modern modem standards, especially when one considers that many kilobytes of software code are usually now found in commerical microprocessor-controlled modem products.

4. SUMMARY

Computationally efficient Recursive-Least-Squares (RLS) procedures have been presented specifically for the adaptive adjustment of the Data-Driven Echo Cancellers (DDECs) that are used in high speed full-duplex data transmission over two-wire telephone lines. The methods have been shown to yield very short learning times for the DDEC while they also are shown to reduce computational requirements simultaneously to levels below those that are required by the most efficient existing RLS (SF) method [12]. During the initialization period, the new numerically stable methods significantly outperform slower-learning stochastic-gradient (LMS) solutions while also requiring no more computational operations than these same LMS solutions.

The new methods can be used with any training sequence over any number of iterations. The new methods are applications of the Fast Transversal (FTF) RLS adaptive filtering algorithms of [13-14]. However, we additionally exploit several special features of the DDEC to dramatically reduce computation below the levels that would have been required for a straightforward use of these FTF algorithms. Several tradeoffs between computation, memory, learning-time and performance have been illustrated. The results of this paper can now be used to design cost-effective high-performance DDEC's for full-duplex data communications with acceptable "start-up".

REFERENCES

[1] V. G. Koll and S. B. Weinstein, "Simulataneous Two-Way Data Transmission Over a Two-Wire Circuit," *Trans. on Com.*, Vol. 21, No. 2, (February 1973), pp. 143-147.

[2] J. J. Werner, "An Inband Data-Driven Echo Canceler for 4800 bps Full-Duplex Data Transmission," *BSTJ*, 1/85.

[3] J.J. Werner, "An Echo-Cancellation-Based 4800 Bit/s Full-Duplex DDD Modem," using Intersymbol Interpolation," *IEEE J. on SAC*, 9/84, pp. 722-730.

[4] D. D. Falconer, K. H. Mueller and S. B. Weinstein, "Echo-Cancellation Techniques for Full-Duplex Data Transmission on Two-Wire Lines," *Proc. NTC'76*, Dallas.

[5] S. B. Weinstein, "A Passband Data-Driven Echo Canceller for Full-Duplex Data Transmission on Two-Wire Circuits," *IEEE Transactions on Communications*, Vol. 25, No. 7, (July 1977), pp. 654-666.

[6] K. H. Mueller, "A New Digital Echo Canceller for Two-Wire Full-Duplex Data Transmission," *IEEE Trans. on Com.*, 9/76, pp. 956-967.

[7] N. A. M. Verhoeckx, *et al.*, "Digital Echo Cancellation for Data Transmission," *IEEE Trans. on ASSP*, 12/79, pp. 768-781.

[8] H. C. Van den Elzen, P. J. Van Gerwen and W. A. M. Snijders, "Echo Cancellation in a Full-Duplex Data Transmission System with Bipolar Encoding, *Proc. NTC'76*, Dallas.

[9] D. Messerschmitt, "Echo Cancellation in Speech and Data Transmission," *IEEE J. on SAC*, 4/84, pp. 283-297.

[10] J. M. Cioffi and J. J. Werner, "The Tap-Drifting Problem in Digitally Implemented Data-Driven Echo Cancellers," *BSTJ*, January 1985.

[11] M. L. Honig, "Echo Cancellation of Voice-Band Data Signals Using Recursive Least Squares and Stochastic Gradient Algorithms," to appear, *IEEE Trans. on Com.*, 1/85, pp. 65-73.

[12] J. Salz, "On the Start-Up Problem in Digital Echo Cancelers," *BSTJ*, 7/83, pp. 1353-1364.

[13] J. M. Cioffi and T. Kailath, "Fast, Recursive-Least-Squares, Transversal Filters for Adaptive Filtering," *IEEE Trans. on ASSP*, 4/84, pp. 304-337.

[14] J. M. Cioffi and T. Kailath, "Windowed Fast Transversal Filters Adaptive Algorithms with Normalization," in *IEEE Trans. on ASSP*, to appear, 1985.

[15] C. W. Farrow, *et al.*, patent pending.

[16] J.M. Cioffi and T. Kailath, "An Efficient RLS Data-Driven Echo Canceller for Fast Initialization of Full-Duplex Data Transmission," to appear, *IEEE Trans. on Communications*, 1985.

[17] A. Milewski, "Periodic Sequences with Optimal Properties for Channel Estimation and Fast Start-Up Equalization," *IBM J. R&D*, 9/83, pp. 426-431.

[18] J. G. Proakis, *Digital Communications* (New York: McGraw-Hill), 1983.

[19] L. Guidoux and B. Peuch, "A Binary Passband Echo Canceller in a 4800 bit/s Two-Wire Duplex Modem," *IEEE J. on SAC*, 9/84, pp. 711-721.

Table 2
SPECIFIC DDEC INITIALIZATION COMPARISON
[illegible]

Algorithm	[illegible]	[illegible]	[illegible]	[illegible]	Start-up Time
[illegible]					
FTF					
Multichannel	352	176	[illegible]	176 at beginning	37ms
Single-channel	176	176	[illegible]	88 begin / 88 middle	37ms
SHF					
Multichannel	352	352	[illegible]	176 at beginning	[illegible]
Single-channel	176	.	[illegible]	88 begin / 88 middle	[illegible]
[illegible]					
FTF					
Multichannel	176	176	[illegible]	0	[illegible]
Single-channel	88	176	[illegible]	88 middle	[illegible]
Exact FTF stable (N not large)	44	176	[illegible]	44 middle	[illegible]

Table 3
DDEC STORAGE COMPARISON FOR TRAINING
(N=22, l=4)

Algorithm	RAM	ROM	ROM(k=22)	ROM(kbyte)[†]
Covariance (Unwindowed) Solutions				
FTF	.	.	.	.
Multichannel	(2l+2)N	2Nk	968	2
Single-channel	(2l+1)N	Nk	484	1
SHF	.	.	.	.
Multichannel	(2l+2)N	.	.	.
Single-channel	(2l+1)N	.	.	.
Prewindowed Solutions				
FTF	.	.	.	.
Multichannel	(2l+2)N	Nk	484	1
Single-channel	(2l+1)N	$\frac{1}{2}$Nk	242	.5
Exact FTF	(2l+1)N	.	.	.
Stable (N not large)	.	.	.	.

END

FILMED

12-85

DTIC